Example 1: The Effects of α, β, and σ on Sample Size

E1-1

	α = 1.0%		**α = 5.0%**	
	β = 15.0%	**β = 20.0%**	**β = 15.0%**	**β = 20.0%**
σ = 1	n = 15	n = 13	n = 9	n = 8
σ = 2	n = 48	n = 43	n = 31	n = 27
σ = 3	n = 105	n = 94	n = 67	n = 57

E1-2
When n = 31, σ = 2, <u>and β = 15.0%</u> the decision-maker is willing to walk away from a dirty site **5%** of the time. This is an **α** or **Type I** error.

E1-3
When n = 31, σ = 2, <u>and α = 5.0%</u> the decision-maker is willing to clean up a clean site **15%** of the time. This is an **β** or **Type II** error.

E1-4
The **standard deviation** (**σ**) has the biggest impact on the number of samples required. Notice as you go down the columns, σ increases at a steady rate from 1 to 2 to 3. The number of samples required, however, does not follow a similar pattern. Instead, a unit change in σ causes an exponential change in the number of samples required.

E1-5
As illustrated in the example above, a small underestimation of σ can lead to a significant underestimation of the number of samples needed to meet the required levels for α and β (the tolerable limits on decision errors). If too few samples are taken and the required error tolerances are not met, the actual error tolerances will be unknown or uncontrolled. In other words, the uncertainty of the study being conducted will not be managed.

Example 2: Judgmental Sampling versus Simple Random Sampling

E2-1

The appropriateness of a sampling plan always depends on the desired use of the data. For the purpose of characterizing the eight stained areas in this example, taking eight judgmental samples (one from each stain) is appropriate. The reason that a judgmental sampling design is appropriate here is that the goal is simply to characterize the eight stains and not to make inferences beyond the specific sites (the eight stains) that were sampled.

E2-2

For the purpose of selling the site, the surface soil within the entire fenced area needs to be characterized. So the question becomes, can the eight judgmental samples be used to characterize the soil within fenced area as a whole? The answer is No. The reasoning is as follows: Each of the eight stained areas had a 100% chance of being sampled and the remaining surface soil with the fenced area had a 0% chance of being sampled. This being the case, the population from which the eight samples were drawn is confined to the eight stains themselves. The rest of the soil within the fenced area is outside the population for this sampling scheme. From a statistical standpoint, it is inappropriate to make inferences to areas outside of the population being sampled.

EPA QA/G-5 (page 20) provides helpful guidance regarding judgmental sampling and inferences:

> "Judgmental sampling does not allow the level of confidence (uncertainty) of the investigation to be accurately quantified. In addition, *judgmental sampling limits the inferences made to the units actually analyzed*[emphasis added], and extrapolation from those units to the overall population from which the units were collected is subject to unknown selection bias."

E2-3

No, for reasons similar to those stated above. The analogous site is outside the population being sampled. Therefore, it would be inappropriate to make inferences to analogous sites, because all samples within the analogous site have 0% chance of being sampled with the current judgmental sampling design.

E2-4

When conducting judgmental sampling, there is no way to identify or control Type I and Type II errors. In other words, the probability of walking away from a dirty site and the probability of cleaning up a clean site are unknown and uncontrolled.

E2-5

As discussed above, judgmental sampling does not allow inferences beyond the specific samples that were taken. In addition, with judgmental sampling there is no control of Type I and Type II errors. The case where a predetermined number of samples are randomly placed within an area is somewhat more complicated. Because the samples are randomly placed, it is correct to call the design a statistical (rather than judgmental) sampling design. However, taking a predetermined number of samples does not allow the Type I and Type II error rates to be known and controlled a priori.

E2-6
No, the Type I and Type II errors were not controlled or managed.

E2-7
No, the conclusions drawn from the eight predetermined samples cannot be extended to analogous sites. The reasoning is similar to that in answer **E2-2**. The analogous site is not part of the population being sampled because all possible samples within the analogous site will not have a chance of being sampled.

E2-8
If $\alpha = 30.0\%$ and $\beta = 30.0\%$, the number of samples required will equal 8. There may be other combinations of α and β that will produce the same result.

E2-9
In this example, the buyer is most heavily impacted by a high α level. If $\alpha = 30\%$, the there is a 30% chance that the buyer will believe that the fenced area is clean when it is actually contaminated. Said another way, there is a 30% chance that the seller will walk away from a dirty site. It is doubtful that the buyer would find this acceptable. The seller is most heavily impacted a high β level. If $\beta = 30\%$, then there is a 30% chance that seller will believe that the fenced area is contaminated when it isn't. Said another way, there is a 30% chance that seller will be required clean up a clean site. Again it is doubtful that the seller would likely find this acceptable. It should be noted, however, that the acceptability of tolerable decision error rates is dependent upon the particular situation.

E2-10
If smaller values for α or β are inserted in the text box, the resultant number of samples will be larger than 8. If larger values for α or β are inserted in the text box, the resultant number of samples will be smaller than 8.

Example 3: Parametric versus Non-parametric Statistics

E3-1
The number of samples required is 48. (where the numerical values of α = 1.0%, β = 25.0%, δ = 1.6, σ = 2.5, and Action Level = 5

E3-2
If a parametric test were performed, it would require 25 samples and the distributional assumptions would need to be tested. If the non-parametric test were performed it would require 48 samples and there would be no need to test the distributional assumptions. Parametric tests have more power to reject the null hypothesis on the basis of fewer samples. So in this case, it would be worthwhile to take 5 additional samples (resulting in 25 + 5 = 30 samples) in order to test the distributional assumptions and possibly be able to use the parametric test. If the distributional assumptions are verified, then the 30 samples already taken are sufficient. If the distributional assumptions are not verified, then the non-parametric test will need to be used and 18 additional samples will need to be taken for a total of 48 samples.

E3-3
The number of samples required with the MARSSIM Sign Test is 60. The number of samples required with One Student t-Test is 31.

Example 4: One Sample Proportion Test

E4-1
The Null Hypothesis for the project is Ho: $P \geq 0.2$ (i.e. the proportion of defective cans in the warehouse is ≥ 20%).

E4-2
The Action Level for this project is 20% of the can inventory are defective.

E4-3
The Action Level illustrated in this example is high because the consequences of leaving the defective cans in place without repackaging has very low risk as the warehouse is not actively used and the consequences of (unnecessary) repackaging is very expensive.

E4-4
The Action Level for this project is a proportion, while the Action Levels normally encountered in soil remediation or D&D projects are concentrations or activity levels.

E4-5
The plant manager would be advised to set the Type I error, α to be 20%, because the consequences of leaving the defective cans in place without repackaging has very low risk.

E4-6
The plant manager would be advised to set the Type II error, β to be as low as possible = 1%, because repackaging is expensive and unnecessary repackaging is wastefully expensive.

E4-7
The minimum number of drums to be inspected under the given conditions is 108.

Example 5: Finding Hot Spots

E5-1
When the **Random Start** option on the **Grid** tab is checked, the number of samples suggested by VSP can vary between 126 and 135 depending on how the **Triangular Grid Type** fits into the one-acre site.

If the **Random Start** option in the **Grid** tab is turned off, about 126 samples will be required to have 95% confidence that a hot spot with a 10-foot radius or larger will be detected.

The smallest circular hot spot that can be detected has a radius of 10.351616 ft and covers an area of 336.640351 ft^2.

E5-2
The required spacing for the triangular grid is 20 ft.

Please note that in VSP, the spacing of the triangular grid is not a function of the number of samples that can be obtained using the **Random Start** option.

E5-3
The total cost of hot spot detection for the entire 10 acre site is 10*$ 63,000 =
$ 630,000 (for 126 samples).

E5-4
For a fixed criteria of 95% confidence to detect a circular 10-ft radius hot spot:

The only options that can be tried are to use a square or rectangular grid other than the selected triangular grid or to minimize the Measurement Cost per Replicate (i.e. offsite analysis cost).

With the **Random Start** option checked, use of a square grid gives a cost of approximately $54,500 and the use of a rectangular grid gives a cost of $109,000 for one 1-acre site. Compare these costs with $63,000 for one 1-acre site that will be expended using a triangular grid

Minimizing or removing the offsite analysis cost and replacing it with a validated lower field screening cost will provide the best return for the investment made by the contractor.

E5-5
For a fixed criteria of 95% confidence to detect a circular 20-ft radius (at least) hot spot and using a triangular grid (with the length of grid size being 40.0 ft) the total cost of hot spot detection for the entire 10 acre site is 10*$ 16,000 = $160,000. For this example, if the hot spot radius is doubled, the total cost to detect is decreased by a factor of ≈ 4.

E5-6

For a fixed criteria of 95% confidence to detect an ellipse having a semi-major axis of 10-ft., shape of 0.6 and using a triangular grid the total cost of hot spot detection for the entire 10 acre site is 10*$ 128,500 = $1,285,000. For this example, if the hot spot shape is changed to the ellipse, the total cost to detect is increased by a factor of ≈ 2.

Example 6: What if I don't know anything about the site?

E6-1
The number of samples required is 305.

E6-2
The number of samples required is 115.

E6-3
The number of samples required is 73.

E6.4
The number of samples required is 57.

E6-5
The number of samples required is 37.

E6-6
The number of samples required is 27.

E6-7
Yes, if historical data is available, the number of samples required may be reduced because a more realistic (and often lower) estimate of the standard deviation is available. In addition, information regarding the shape of the distribution may also be obtained.